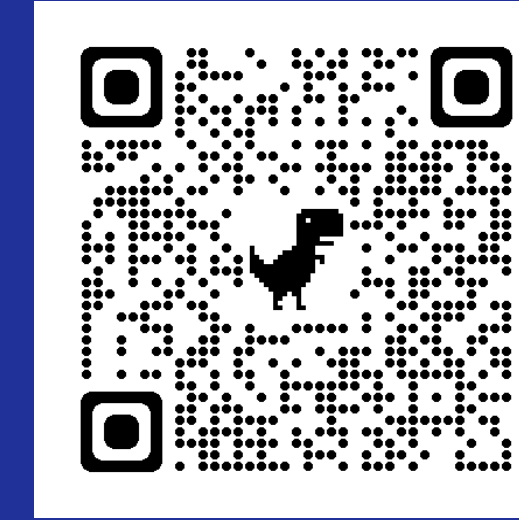


Smart Vision Systems Laboratory (SVSL)



Senem Velipasalar, Ph.D., Professor and Director, SVSL
Dept. of Electrical Engineering and Computer Science
Associate Editor, IEEE Transactions on Image Processing



Mission/Purpose

- To develop advanced machine learning algorithms, and apply them to real-world data captured for different applications from a variety of sensors, including but not limited to visible-range and thermal cameras, microphones, accelerometers and fNIRS devices
- To collaborate across disciplines and work with researchers in academia and industry to address real-world problems
- To train students to become next generation engineers and scientists in the rapidly expanding areas of machine learning, embedded systems and computer vision

Scope/Current Research

The research conducted in the SVSL is mainly focused on computer vision, mobile cameras and designing advanced machine learning algorithms for different applications

MicroCam Overview

MicroCam consists of low-power, AI-based, IoT platforms

- Each device has multi-modal sensors and can process motion, audio and video data
- Send binary occupancy result to the "lead platform"
- Stand-alone solution
- Battery-powered for easy self-commissioning
- All processing is performed on platforms. No use of external or cloud computing
- Preserves privacy
- Only the binary occupancy state is shared with the lead platform
- Preliminary results show promise on images that are modified so that occupants are not discernable

Product Benefits and Performance

- Can operate under daylight, low light and no-light conditions
- Can detect stationary occupants
- Can differentiate people from pets and other sources of motion (e.g., robot vacuum, fans, blowing curtains, etc.)
- Preserves privacy

Real-world Testing

- Covering difficult scenarios, including people lying down, people in chairs, scenes with cats, very low light and no-light conditions
- The overall accuracy of the real-time, integrated tests, lasting over 111 hours, is 92.0% with no false positives
- Additional 269 hours of testing at apartments with an overall accuracy of 99.3%
- 94-97% accuracy on privacy-protection images
- Surveillance and security monitoring
- Aging-in-place and Nursing Homes: Activity Monitoring and Fall Detection
- Smart homes
- Additional capabilities can be added to the system including fine-grained activity classification, determining duration of activities, and object interactions, and detecting other activities or events of interest

MicroCam: A Low-Power and Privacy Preserving Multi-modal Sensor Platform for Occupancy Detection for HVAC Control and Surveillance

Why Discard if You Can Recycle? A Recycling Max Pooling (RMP) Module for 3D Point Cloud Analysis

Our proposed RMP module first obtains F_1, F_2, \dots, F_n by recycling the discarded points, and then refines F_i by designing a Hierarchical Loss Function. This loss function incorporates classification loss and refinement loss.

Classification Results on the ModelNet40 Dataset

Model Name	Highest Acc.	Smoothed Acc.
PointNet++	90.32%	88.37%
PointNet++ (w/)	90.32%	88.37%
GCNet	88.37%	86.39%
GCNet (w/)	88.37%	86.39%
DPPA	91.46%	89.36%
DPPA (w/)	91.46%	89.36%
ConvNext	92.82%	90.59%
ConvNext (w/)	92.82%	90.59%

Classification Results on the ScanObjectNN Dataset

Model Name	Highest Acc.	Smoothed Acc.
PointNet++	89.01%	87.11%
PointNet++ (w/)	89.01%	87.11%
GCNet	87.11%	85.28%
GCNet (w/)	87.11%	85.28%
DPPA	90.29%	88.37%
DPPA (w/)	90.29%	88.37%
ConvNext	93.42%	91.46%
ConvNext (w/)	93.42%	91.46%

Aerial Building Inspection with UAVs equipped by Infrared Cameras

PROBLEM

- Autonomous detection of thermal anomalies from thermal images captured from UAVs
- Have designed machine learning algorithms and applied them to data captured from IR cameras
- 85% time savings** in envelope building energy audit
- 90% cost reduction** compared to typical envelope energy audit
- Safer & More Accurate**

40% energy savings in envelope energy audit

50% cost reduction compared to typical envelope energy audit

Thermal Anomalies

Anomaly Classification: Thermal Budget: 100.00

Point Cloud Classification and Segmentation

Problem Definition

Proposed Method

Motivation

- Large-scale scene changes and fast camera motion make egocentric action recognition (EAR) a challenging problem in computer vision. Egocentric view captures objects as well as the interactions between objects and the subject.
- Although transformer-based models have been introduced as powerful tools in video understanding and achieved state-of-the-art performance, only a handful of them specifically considers the properties of the first-person view.
- We focus on improving the performance of video transformers for EAR. More specifically, we propose a novel framework, referred to as the EgoViT, which carefully takes the special properties of egocentric data into account, and can be integrated with different video transformers.

EgoViT: Pyramid Video Transformer with Dynamic Class Token for Egocentric Action Recognition

Limitations of Existing Methods:

- Point clouds: vulnerable to occlusions
- Depth images: can increase robustness to occlusions
- Pooling operations discard useful information
- The correlation between support and queries is not fully utilized

Solutions:

- Combine two data modalities to enrich features
- Pyramid pooling: gather global & local information
- Proposed Support-Query Mutual Attention module

Model architecture for few-shot 3D point cloud classification

	field 1	field 2	Average
5-way 1-shot w/o sqMA	58.18	64.53	66.87
5-way 1-shot w/ sqMA	61.09	66.29	68.39
5-way 5-shot w/o sqMA	72.84	75.81	81.29
5-way 5-shot w/ sqMA	74.00	76.51	83.02

Cross-Modality Few-Shot Point Cloud Classification

Motivation

- Large-scale scene changes and fast camera motion make egocentric action recognition (EAR) a challenging problem in computer vision. Egocentric view captures objects as well as the interactions between objects and the subject.
- Although transformer-based models have been introduced as powerful tools in video understanding and achieved state-of-the-art performance, only a handful of them specifically considers the properties of the first-person view.
- We focus on improving the performance of video transformers for EAR. More specifically, we propose a novel framework, referred to as the EgoViT, which carefully takes the special properties of egocentric data into account, and can be integrated with different video transformers.

EgoViT: Pyramid Video Transformer with Dynamic Class Token for Egocentric Action Recognition

Vision-based Gait Recognition

Current Research Cont'd

Deep neural networks have been widely used in different application areas, such as autonomous driving, face recognition, image classification etc.

Yet, they have been shown to be vulnerable to adversarial attacks and carefully crafted adversarial examples.

How can we detect adversarial examples?

We have proposed two methods for detecting adversarial examples: one is for autonomous driving scenarios, and the other is for detecting adversarial examples.

There have been works focusing on making a neural network more robust against adversarial attacks by hardening the network itself.

Yet, attacks have been more advanced and can easily evade the defense mechanisms. Thus, attention has been paid to detecting adversarial examples.

Adversarial Example Detection

Classification Model

SLA via Rest-Task Transf.

Transparency Study

Cognitive Workload Classification

Performance Drop

Position Index

Projects

- MicroCam: A Low Power and Privacy Preserving Multi-Modal Platform for Occupancy Detection**, in partnership with SRI International, sponsored by Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy and NYSEERDA
- Aerial Intelligence for Retrofit Building Energy Modeling (AirBEM)**, in partnership with Georgia Tech and MIT, sponsored by Department of Energy - Buildings Energy Efficiency Frontiers & Innovation Technologies (BENEFIT)
- Deep Insight - Deep-Net Driven Approach to Estimate Driver State in Naturalistic Data**, in partnership with Iowa State, sponsored by the Department of Transportation
- Improved Cross-Subject Cognitive and Emotional State Classification Using Functional Near-Infrared Spectroscopy Data for Deep Learning**, sponsored by National Science Foundation

Major Contributions/Output

- MicroCam prototype for HVAC control and surveillance
- Patent: S. Velipasalar, A. Almagambetov, M. Casares, **"Automatic Detection by a Wearable Camera,"** 9,571,723, issued on February 14, 2017
- Edited Books:
 - Embedded, Cyber-Physical, and IoT Systems
 - Distributed Embedded Smart Cameras

- Selected publications at https://ecs.syr.edu/faculty/velipasalar/publications_velipasalar.html

Team Members

- Dr. Senem Velipasalar, Professor
- Ms. Chenbin Pan, Ph.D. candidate (Sep. 2019 to present)
- Mr. Jiyang Wang, Ph.D. candidate (Sep. 2019 to present)
- Mr. Weiheng Chai, Ph.D. candidate (Sep. 2020 to present)
- Mr. Jiajing Chen, Ph.D. candidate (Sep. 2020 to present)
- Ms. Minmin Yang, Ph.D. student (Sep. 2021 to present)
- Ms. Huantao Ren, Ph.D. student (Sep. 2022 to present)



Sponsors

- Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy
- New York State Energy Research & Development Authority (NYSEERDA)
- U.S. Department of Energy
- U.S. Department of Transportation
- National Science Foundation
- Syracuse University, CUSE grant



vision.syr.edu

